Testing a Model of Consumer Vehicle Purchases

Draft

United States
Environmental Protection
Agency

Testing a Model of Consumer Vehicle Purchases

Draft

Gloria Helfand*[a]
Changzheng Liu[b]
Marie Donahue[c]
Jacqueline Doremus[c]
Ari Kahan[a]
Michael Shelby[a]

*Corresponding author:
[a]Office of Transportation and Air Quality
U.S. Environmental Protection Agency
2000 Traverwood Drive
Ann Arbor, MI 48105 US
helfand.gloria@epa.gov

[b]Oak Ridge National Laboratory
PO Box 2008 MS6472
Oak Ridge, TN 37831-6472 US

[c]Oak Ridge Institute for Sicience and Education Program
Office of Transportation and Air Quality
U.S. Environmental Protection Agency
2000 Traverwood Drive
Ann Arbor, MI 48105 US

United States Environmental Protection Agency

EPA-420-D-15-011
December 2015

TESTING A MODEL OF CONSUMER VEHICLE PURCHASES

ABSTRACT

Consumer vehicle choice models have been estimated and used for a wide variety of policy simulations. Infrequently, though, have predicted responses from these models been tested against actual market outcomes. This paper presents a validation exercise for a model developed for the U.S. Environmental Protection Agency, intended to estimate the impacts of changes in vehicle prices and fuel economy due to changes in vehicle greenhouse gas emissions standards. The model is a nested logit with a representative consumer and 5 levels, calibrated to vehicle purchases in model year (MY) 2008. First, we review the model's response to a simple policy scenario, to explore effects of different parameter values on the outcomes of that scenario; we find that the model is not particularly sensitive to key parameters. Next, vehicle changes between MY 2008 and 2010 are used to make predictions, and those predictions are compared to actual outcomes in MY 2010; the model, designed to examine changes due only to price and fuel economy, did not do as well in predicting sales impacts as assuming that market shares were the same as in MY 2008, during this time of significant economic change. These exercises raise questions about how to validate a model intended for comparative static analysis in a dynamic world.

Keywords: vehicle demand; consumer vehicle choice modeling; validation; discrete choice modeling

Testing a Model of Consumer Vehicle Purchases

1 Introduction

How well can a model predict which cars people will buy? Modeling purchase patterns of consumer vehicles matters because of the importance of the auto sector to the U.S. economy, and because of the contributions of vehicles to air pollution, including greenhouse gas (GHG) emissions. Automakers have clear incentives to estimate consumer vehicle desires well, especially when it may be difficult to change production plans quickly in response to market signals. Many public policies, such as the federal Car Allowance Rebate System ("Cash for Clunkers") in 2009 or California's Zero Emissions Vehicle Program, have explicit goals to affect what vehicles people buy. Other policies, such as greenhouse gas or fuel economy standards, may indirectly affect how many and which vehicles people buy. Measuring the effects of these programs on vehicle sales would provide greater insights into the impacts of these programs on the environment, auto producers, and the public.

Many researchers have developed models designed to estimate vehicle purchases. The models, commonly econometrically estimated, are often used for prospective simulation purposes. Almost unstudied, though, is the effectiveness of these models in predicting market responses to changed circumstances. Researchers have rarely used their models to examine situations where model results could be compared to actual market outcomes.

It seems evident that the utility of these models for estimating policy impacts should depend on their effectiveness in predicting those impacts. It is also to be expected that the success of a model depends on the purposes for which it was designed: a model well designed for examining the effects of GHG standards, for instance, may not perform well in estimating the effects of demographic shifts on buying patterns.

The U.S. Environmental Protection Agency (EPA) has been exploring the use of vehicle choice models in analyzing the impacts of vehicle GHG/fuel economy regulations, and recently commissioned a consumer vehicle choice model for potential use in its analysis of the impacts of vehicle greenhouse gas regulations on the U.S. auto market (Greene and Liu 2012). This paper presents results of a validation exercise for this model. This validation exercise contains two parts. In the first, we increase the fuel economy of all vehicles by 20 percent, and then examine the effects of changing key parameters on modeling results; we find that the model's results are not especially sensitive to changes in these parameters. Next, with the model calibrated to sales of model year (MY) 2008 vehicles, we use the actual changes in vehicle prices and fuel economy for MY 2010 vehicles to predict the sales of MY 2010 vehicles; we then compare the predicted sales to actual sales. Here, we find that the model, designed to examine changes due only to price and fuel economy, did not do well as in predicting sales impacts as assuming that market shares were the same as in MY 2008, during this time of significant economic change. Perhaps a key finding of this exercise is that model validation is both very important and potentially very difficult to assess.

2 Background

The magnitude of the auto industry in the U.S. economy and the importance of its role in international trade and environmental protection have led to dozens of articles that analyze the impacts of various factors and policies on consumer vehicle purchases. For instance, Goldberg (1998), Whitefoot and Skerlos (2011), and Jacobsen (2013) examine the effects of fuel economy standards; Greene (2009) considers feebates; Train and Winston (2007) test the competitiveness of the U.S. auto industry; and Brownstone et al. (1996) model the market acceptability of alternative-fuel vehicles. Helfand and Wolverton (2011) review this literature, though the literature continues to expand (e.g., Bento et al. 2012; Allcott 2013).

In most of these papers, the quality of the model is based on the econometric analysis: if the analysis meets theoretical and statistical requirements, and the results include expected and statistically significant coefficients on variables, then the model is considered suitable for policy analysis. Researchers commonly use their models for simulation of counter-factual situations based on the best estimates of the baseline situation. For instance, Goldberg (1998), Austin and Dinan (2005), and Jacobsen (2013) all assess the relative merits of a gasoline tax vs. fuel economy standards. Goldberg found that gasoline prices would have to double to get the same effect on fuel consumption as fuel economy standards, due to a low estimate of responsiveness to operating costs; both Austin and Dinan and Jacobsen, on the other hand, find a gasoline tax to be much more efficient.[1]

Despite their widespread use for policy simulation, these models have typically not been validated for their ability to predict vehicles sales in response to new circumstances. That is, rarely have their predictions been tested against real-world outcomes, to see if they can in fact predict out of sample. In other disciplines, this cross-checking of modeling to actual outcomes is a significant aspect of the research agenda. For instance, experimental economists often test hypothetical scenarios against actual market behavior (e.g., Landry and List 2007) to examine

[1] These models are not directly comparable. Unlike Goldberg's model, Austin and Dinan's and Jacobsen's models take into account the used vehicle fleet. Because a gasoline tax affects existing vehicles as well as new vehicles, it saves fuel across the fleet. In contrast, a fuel economy standard affects only new vehicles.

the validity of stated preference studies. For evaluating air quality modeling simulations, the U.S. Environmental Protection Agency has developed the Atmospheric Model Evaluation Tool specifically to compare predictions about meteorology and air quality against outcomes (U.S. Environmental Protection Agency 2014).

One exception in the vehicle modeling literature is Pakes et al. (1993), who (as summarized in Berry et al. 1995):

> . . . used our model's estimates to predict the effect of the 1973 gas price hike on the average MPG [miles per gallon] of new cars sold in subsequent years. We found that our model predicted 1974 and 1975 average MPG almost exactly. . . . However, by 1976 new small fuel efficient models began to be introduced and our predictions, based on fixed characteristics, became markedly worse and deteriorated further over time.

Another exception is Haaf et al. (2013), who use data from MY 2004-6 vehicles to estimate a number of different econometric models, and test their predictions against MY 2007 and 2010 vehicle sales. The models had an average error of 0.24 percent compared to a mean vehicle share of 0.42 percent: in other words, "the models we construct are fairly poor predictors of future shares." They find that a "static" model – that is, one that assumes constant market shares -- outperformed their estimated models for MY 2007, while the attribute-based models predicted better for MY 2010. Finally, they caution that "some of the models with the best predictive accuracy have coefficients with unexpected signs – likely biased due to correlation with unobserved attributes."

Finally, Raynaert (2014) develops a structural model of vehicle supply and demand in Europe, using data from 1998-2007; he then compares sales-weighted aggregate predictions from the model for MY 2011 to actual outcomes. He finds close agreement: in a period where actual emissions dropped 14 percent, his estimates for emissions differed from the observed values by 2.3 percent. Weight, footprint, and the share of diesel also had discrepancies of 3 percent or less; price/income and horsepower differed by under 10 percent. He does not provide information about more disaggregated results.

The paucity of research assessing the performance of vehicle choice models, along with these papers, suggests that the predictive ability of consumer vehicle choice models is not yet proven.[2] At the same time, analysis of the impacts of policies on vehicle sales and class mix require some prediction, whether a static approach of constant market shares, or a more sophisticated analysis that accounts for future vehicle or market characteristics. This paper adds to that literature by performing a validation exercise on a consumer vehicle choice model developed for the U.S. Environmental Protection Agency (EPA). This model is designed for very specific policy simulations: to examine the effects of changes in fuel economy and vehicle price on U.S. vehicle purchase patterns in response to GHG standards. The fact that it was designed for static policy simulations rather than for forecasting raises additional issues for model validation.

3 The Model[3]

Greene and Liu (2012) developed the vehicle choice model used here for EPA specifically to predict changes in total sales and fleet mix associated with GHG/fuel economy standards. As will be discussed further, it is intended to compare a specified fleet with and

---

[2] In addition, Knittel and Metaxoglou (2014) find that the estimation method Berry et al. (1995) and Raynaert (2014) used is sensitive to start values and optimization algorithms, with results varying by substantial margins.
[3] This section draws heavily from Greene and Liu (2012).

without fuel economy standards;[4] it is thus a static model, not intended to account for changes in macroeconomic or demographic conditions.

It is a nested logit with a representative consumer and 5 layers, as described in Figure 1 and Table 1. The first layer constitutes the buy/don't buy decision. Next it distinguishes between passenger vehicles, cargo vehicles, and ultra-prestige vehicles. In the model, sport-utility vehicles and minivans are included as passenger vehicles, although many of these vehicles are considered light-duty trucks for regulatory purposes. Consumers commonly consider these to be passenger vehicles; it is more likely, for instance, that people consider an SUV to be a substitute for a large or midsize car than for a pickup truck. Because the model is meant to reflect consumer decision processes, it was considered appropriate to nest SUVs and minivans as passenger vehicles rather than cargo vehicles. Ultra-prestige vehicles are defined as those with price exceeding $75,000.

[4] EPA regulates GHG emissions from vehicles; the Department of Transportation regulates vehicle fuel economy. Because the primary way to reduce GHG emissions is to improve fuel economy, the agencies harmonized their regulations (U.S. EPA and Department of Transportation 2010, 2012). The model uses fuel economy rather than GHG emissions, because fuel economy is much more salient an attribute to vehicle buyers.

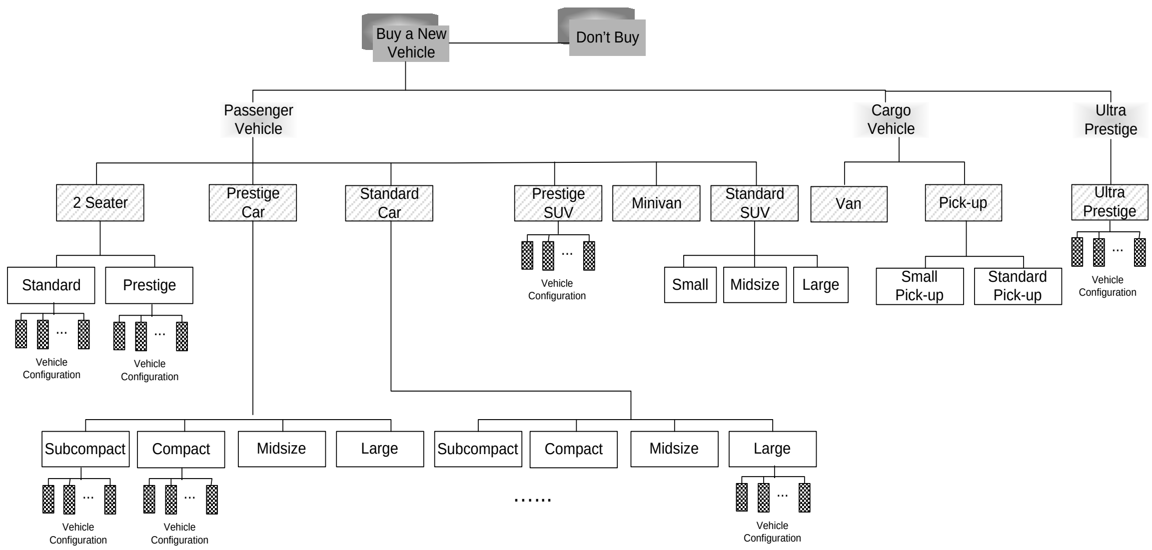

Figure 1: Nested Multinomial Logit Structure of Consumer Choice Model
Note: "Standard" is synonymous with "Non-Prestige"

Table 1: Vehicle Class Definition in the Consumer Vehicle Choice Model

Model Class	Corresponding EPA Class
1. Prestige[1] Two-Seaters	Two Seaters
2. Prestige Subcompact Cars	Subcompact Cars, Minicompact Cars
3. Prestige Compact Cars and Small Station Wagons	Compact cars, Small Station Wagons
4. Prestige Midsize Cars and Station Wagons	Midsize Cars, Midsize Station Wagons
5. Prestige Large Cars	Large Cars
6. Two-Seater	Two Seaters
7. Subcompact Cars	Subcompact Cars, Minicompact Cars
8. Compact Cars and Small Station Wagons	Compact Cars, Small Station Wagons
9. Midsize Cars and Station Wagons	Midsize Cars, Midsize Station Wagons
10. Large Cars	Large Cars
11. Prestige SUVs	SUVs
12. Small[2] SUVs	SUVs
13. Midsize SUVs	SUVs
14. large SUVs	SUVs
15. MiniVans	MiniVans
16. Cargo/Large Passenger Vans	Cargo Vans, Passenger Vans
17. Small Pickup Trucks	Small Pickup Trucks
18. Standard Pickup Trucks	Standard Pickup Trucks

19. Ultra Prestige Vehicles[3]	See the definition (note 4) below

Notes:

(1) Prestige and non-prestige classes are defined by vehicle price: the prestige are vehicles whose prices are higher than or equal to unweighted average price in the corresponding EPA class, and vice versa for non-prestige vehicles. E.g., Prestige Two-Seater class is the set of relatively expensive vehicle configurations in EPA class of two seaters with prices higher than or equal to the unweighted average price of EPA two seaters.

(2) Non-prestige SUVs are divided into small, midsize and large SUVs by vehicle's footprint (small: footprint <43; midsize: 43<=footprint<46; large: footprint>=46)

(3) Ultra Prestige class is defined as the set of vehicles whose prices are higher than or equal to $75,000.

The model then separates passenger vehicles into Two Seaters, Prestige Cars, Standard Cars, Prestige SUVs, Standard SUVs, and Minivans (with prestige also determined by price), and cargo vehicles into Pickup Trucks and Vans. The next level continues the division into classes, and the final level consists of individual vehicles. The model is calibrated to sales by individual vehicle type in a base year through use of each vehicle's price and fuel economy. Fuel savings for a vehicle are calculated as the present value, for a user-defined period of time (the "payback period"), of fuel expenditures, based on the vehicle's mpg, vehicle miles traveled, and fuel prices. The price and fuel savings are used to estimate an effective price; when that effective price is combined with the price slope for that vehicle's nest, the constant term is the value that results in matching the initial sales volume for that vehicle.

Vehicle sales are predicted to change in response to changes in net vehicle price, where the change in net vehicle price is calculated as the increase in vehicle cost associated with technologies to reduce GHGs, less a discounted share of the future fuel savings associated with those technologies. Greene (2010) found highly varied estimates in the literature of consumer willingness to pay (WTP) for additional fuel economy in the vehicle purchase decision, with a number of studies showing WTP less than the expected value of future fuel savings, and some others showing overvaluation. The model allows a user to choose the number of years of expected fuel savings that vehicle buyers are believed to consider in their purchase decisions, as well as the future fuel prices and discount rate they might use for those calculations.

The model is designed to interact with EPA's technology-cost model, the Optimization Model for reducing Greenhouse Gases from Automobiles (OMEGA), which seeks cost-effective combinations of technologies to achieve GHG standards (U.S. EPA 2012). Iteration between the model and OMEGA can be used to estimate whether sufficient technology is added to vehicles to bring fleets into compliance with standards, after modeled consumer responses are taken into account.

The demand elasticities in the model for each vehicle nest are not estimated from an original data set, but rather are based on reviewing estimates in the literature (Greene and Liu 2012, Table 4). This approach has advantages and disadvantages. It allows for synthesis of the results from multiple analyses, and professional judgment about whether the values are appropriate. It also can be viewed as combining results from different studies, where the differences in the studies may have implications for the value. Table 2 provides the elasticities used in the analysis.

Table 2: Default Elasticities

Level 4	Choice of Make, Model, Engine Transmission Configuration within a Class				Level 3	Choice Among 19 Vehicle Classes within Vehicle Type		
	Name	Elasti-city	Parent Type			Name	Elasti-city	Parent Category
1	Prestige Two-Seater	-3.8	Two-Seater		1	Two-Seater	-1.3	Passenger
2	Prestige Subcompact	-3.5	Prestige Car		2	Prestige Car	-2.2	Passenger
3	Prestige Compact and Small Station Wagon	-3.5	Prestige Car		3	Standard Car	-3	Passenger
4	Prestige Midsize Car and Station Wagon	-3.6	Prestige Car		4	Prestige SUV		Passenger

5	Prestige Large	-3.5	Prestige Car		5	Standard SUV	-2.7	Passenger
6	Two-Seater	-3.5	Two-Seater		6	Minivan		Passenger
7	Subcompact	-5	Standard Car		7	Cargo Van		Cargo
8	Compact and Small Station Wagon	-5	Standard Car		8	Cargo Pickup	-2	Cargo
9	Midsize Car and Station Wagon	-5	Standard Car		9	Ultra Prestige		Ultra Prestige
10	Large Car	-5	Standard Car					
11	Prestige SUV	-3.7	Prestige SUV		Level 2	Choice of Vehicle Type within Passenger or Cargo Categories		
12	Small SUV	-4.9	Standard SUV			Name	Elasticity	Parent Node
13	Midsize SUV	-5.1	Standard SUV		1	Passenger	-1.1	Buy
14	Large SUV	-5.1	Standard SUV		2	Cargo	-0.7	Buy
15	Minivan	-4.9	Minivan		3	Ultra Prestige		Buy
16	Cargo / large passenger van	-5.1	Cargo Van					
17	Cargo Pickup Small	-5.1	Cargo Pickup		Level 1	Choice of Passenger, Cargo or Ultra Prestige Vehicle		
18	Cargo Pickup Standard	-5.1	Cargo Pickup			Name	Elasticity	Parent Node
19	Ultra Prestige	-3.9	Ultra Prestige			Buy	-0.7	Root
						No Buy		Root

A few limitations of the model are identifiable even before any simulations are run. Some of these limitations arise from the model being designed to be calibrated to an existing fleet and then to estimate deviations from that initial calibration. The model thus does not account for macroeconomic shocks that might affect either total sales or changes in fleet mix independent of GHG standards, the introduction or departure of vehicles in the fleet, changes in consumer preferences, or manufacturer changes in other vehicle characteristics (such as performance or appearance). For the purposes for which the model was built, these are not limitations. The model was designed for static, same-year analysis of the effects on vehicle sales of adding fuel-saving technologies and their costs; that is, it was intended to compare vehicle sales with and without fuel-saving technologies and additional costs for a single fleet of vehicles. In principle, then, changes in the economy, demographics, or the fleet over time should not affect the ability of the model to predict, because it is predicting against a static counter-factual. However, the baseline of no standards and the counter-factual of meeting the standards do not, in reality, exist in the same year. Instead, the model will here be tested for its ability to predict between two model years. As will be discussed further, the years for which we currently have data involve the beginning and the depths of the Great Recession, whose effects may swamp any predictive abilities of the model, and the Cash for Clunkers program in 2009 that may have pulled sales forward from 2010. This limitation is therefore an issue for this method of testing the model.

Other limitations are associated with the use of nested logit. For instance, as Train (2009) notes, "only differences in utility matter." As a result, an equal change in prices (e.g., $1000) for all vehicles in the same nest would lead to no reallocation of market shares among vehicles in that nest, although a $1000 change has a much bigger relative impact on the price of an inexpensive car than that of a more expensive car. (The price increase would change total sales and market shares across nests.) The nested logit also puts restrictions on demand elasticities for the nests: responsiveness to price must be highest at the individual-vehicle level, and decrease at each higher nest. The model includes a validation step to ensure that these elasticity restrictions are achieved. Finally, within nests, logit exhibits "independence of irrelevant attributes" (IIA): the ratio of probabilities (or market shares, in this model) of two options does not vary if a third option is added to the mix. As a result, an increase in the market share of one alternative within a nest draws proportionately from all other alternatives (Train 2009, Chapter 3). Across nests, IIA does not hold. It is thus important for the nests within a nested logit to contain vehicles that are close substitutes for each other, so that this substitution pattern is a reasonable approximation.

The model testing consists of two parts. The first uses a hypothetical 20 percent increase in fuel economy for all vehicles to examine the sensitivity of the model to changes in key parameters, including payback period for fuel economy, discount rate, elasticities, and start values. The second involves calibrating the model to MY 2008 vehicles and then using the changes in fuel economy and price between 2008 and 2010 as inputs to review the ability of the model to predict changes in vehicle sales.

4 Data

Data requirements for the model include the vehicle's price, fuel economy, and sales, as well as the new fuel economy and the change in price. These data come from market data assembled by EPA and the Department of Transportation for their analysis of GHG standards for MYs 2017-25 (U.S. Environmental Protection Agency and Department of Transportation 2012) for both MY 2008 and MY 2010 vehicles. Both datasets contain over 1000 unique vehicles.[5]

The sensitivity analyses use only the MY 2008 data. Both price and fuel economy enter into the calculation of net price that the model uses to estimate sales changes; for hypothetical effects of the model, where the goal is to simulate a relatively arbitrary change in net price, it is

[5] For example, there are 20 different versions of the Chevrolet Silverado in the 2008 data, each unique based on engine, footprint, fuel economy, baseline sales, and other attributes.

not necessary to change both price and fuel economy. The policy simulations therefore involve a 20 percent increase in fuel economy to all vehicles, with no increase in price.[6] In essence, the policy scenario is a reduction in the net price of all vehicles. The net price reduction is greater in absolute terms for vehicles with lower fuel economy, because a 20 percent increase in miles per gallon for, e.g., a vehicle that gets 10 mpg results in a much greater reduction in fuel consumption than a 20 percent increase for a more efficient vehicle (Larrick and Sol 2008).[7] The simulation analyses use the entire MY 2008 vehicle fleet.

An additional needed set of parameters consists of fuel prices, used for the calculation of fuel savings over the period that a vehicle buyer considers in the purchase decision (here called the "payback period"). The sensitivity analyses use fuel prices as projected in the 2008 Annual Energy Outlook (Energy Information Administration 2008). The calculation of fuel savings also uses the schedule of vehicle miles traveled used in U.S. Environmental Protection Agency and Department of Transportation, 2012.

Both the MY 2008 and 2010 datasets are needed for the prediction exercise. In this case, the model takes as input the baseline price and fuel economy of each vehicle in MY 2008, and then uses any change in price and fuel economy between MY 2008 and 2010 to predict sales in MY 2010. Therefore, each MY 2008 vehicle needed to be matched with its MY 2010 counterpart. This matching is not straightforward. Vehicles enter and exit the market between any two model years; indeed, Saab dropped out of the market entirely during this time. This

---

[6] The model calibrates itself to the base year data, so that, if price and fuel economy do not change, the model returns the initial sales, regardless of the values of other parameters. A policy scenario is necessary to produce changes in vehicle sales.

[7] This "mpg illusion" arises because fuel consumption is inversely related to mpg. For a vehicle that drives 15,000 miles per year, switching from a 10-mpg vehicle to a 12-mpg vehicle saves 250 gallons per year; switching from a 30-mpg vehicle to a 36-mpg vehicle saves 83 gallons per year.

paper uses two methods to address this problem. In the first, aggregation by vehicles, multiple trim levels (for instance, two-door vs. four-door versions of a model) of each vehicle are combined through sales-weighting. This approach allows matching of most of the individual vehicle models. In the second, aggregation by class, all vehicles are aggregated, by manufacturer, to the classes of the vehicle choice model (see Table 1 for those classes). In this study, any remaining unmatched vehicles are dropped from the analysis. Table 3 and Table 4 provide the summary statistics for these two methods compared to the whole fleets. Both cases permit matching of over 90% of the vehicles sold in either model year, though aggregating by class allows for representation of somewhat more vehicles.

Table 3: Summary Statistics of Baseline and Aggregated Fleets

	2008			2010		
	Baseline	Fleet Aggregated by Vehicle	Fleet Aggregated by Class	Baseline	Fleet Aggregated by Vehicle	Fleet Aggregated by Class
Total number of unique vehicles	1302	524*	145**	1171	524*	145**
Total vehicle sales	13,851,761	12,976,766	13,573,775	11,190,180	10,199,188	10,648,872
% Total vehicle sales captured in the final matching process	--	94%	98%	--	91%	95%

*108 unmatched vehicles include manufacturers or vehicles manufactured in one year but not in the other. These are dropped in the analyses that follow.
**Two manufacturers (Spyker/Saab, Tesla) had sales in MY 2008 but not MY 2010. In 36 occasions, a manufacturer had sales in a vehicle class in one year but not in the other. These are dropped in the analyses that follow.

Table 4: Additional Summary Statistics

	MY 2008 Actual	MY 2008 aggr. by vehicle	MY 2008 aggr. by class	MY 2010 Actual	MY 2010 aggr. by vehicle	MY 2010 aggr. by class
Total sales (millions)	13.9	13.0	13.6	11.2	10.2	10.6
Weighted avg. price	$27,873	$27,702	$27,850	$26,767	$26,624	$26,861
Minimum price	$11,783	$11,783	$13,646	$9,970	$11,923	$12,816
Maximum price	$1.7M	$1.7M	$254,533	$1.7M	$1.7M	$213,672
Weighted avg. fuel economy	26.2	26.3	25.7	28.4	28.3	27.5
Min fuel economy	12.0	12.0	15.2	12.0	12.0	14.1
Max fuel economy	65.8	65.8	49.5	70.8	70.8	49.1
Share passenger	86.3%	85.7%	86.0%	87.8%	86.8%	87.2%
Share cargo	12.8%	13.4%	13.1%	11.6%	12.7%	12.1%
Share ultra-prestige	0.9%	0.9%	0.9%	0.7%	0.5%	0.7%

Table 4 shows that the two forms of aggregation lead to small differences in fleet characteristics. Aggregating by class matches the full fleet slightly better on weighted average price, but aggregating by vehicle matches slightly better on average fuel economy. Differences in shares among passenger, cargo, and ultra-prestige vehicles are less than one percent in all cases.

5 Sensitivity Analyses

The baseline parameterization of the consumer choice model used the elasticities in Table 2; an assumption that consumers would consider 5 years of fuel savings in their purchases (i.e., a five-year payback period); and a discount rate of 3 percent for calculating future fuel savings. The policy experiment is an across-the-board 20 percent increase in fuel economy, about equivalent in total sales impacts to a reduction in prices for all vehicles of 6.5 percent.[8]

As Table 5 shows, increasing the fuel economy of all vehicles by 20 percent has the expected effect of increasing vehicle sales, from 13.85 million to 14.55 million (5 percent). While sales increase for all vehicle classes, the largest absolute increases in sales occur for Cargo Pickup Standard, Small SUV, and Prestige SUV, which each increase by over 100,000 vehicles, roughly 8 percent in all three cases.[9] The smallest percentage increases, 2.5% or less, were for Subcompacts, Compacts, Small Cargo Pickups, and Two-Seaters. This pattern perhaps reflects the model's use of expected future fuel savings in the net price calculation. The classes with the greatest sales gains had initial fuel economy that averaged between 19 and 23 mpg; the classes with the smallest increases had average initial fuel economy over 30 mpg, except for Small Cargo Pickups (23.5 mpg). As discussed above, the absolute reductions in net prices for the less efficient vehicles were greater than that for the more efficient vehicles, and the model finds greater sales increases for those less efficient vehicles. As a result of the change in sales mix,

[8] Changing price by a uniform percentage leads to different sales mix than changing fuel economy by a uniform percentage, because price and fuel economy are not perfectly correlated.
[9] Others with large percentage increases were also generally large or prestige vehicles, including Prestige Two-Seaters, Large Cars, Minivans, Cargo and Large Passenger Vans, but had much smaller total sales.

fleet average fuel economy is predicted to increase from 26.2 mpg to 31.2 mpg, slightly less than the 20 percent increase applied to all vehicles.

Table 5: Effects on MY 2008 Fleet of Increasing Fuel Economy by 20 Percent

	Initial Sales	Initial Fuel Economy	Final Sales	Percent Change in Sales	Final Fuel Economy
Prestige Two-Seater	75,467	24.8	79,806	5.6%	29.8
Prestige Subcompact	303,812	26.8	317,930	4.5%	32.1
Prestige Compact and Small Station Wagon	389,652	27.4	406,431	4.2%	32.8
Prestige Midsize Car and Station Wagon	587,330	26.2	616,536	4.9%	31.4
Prestige Large	987,537	27.3	1,030,415	4.2%	32.7
Two-Seater	64,730	34.0	66,211	2.3%	40.3
Subcompact	952,113	34.7	970,570	1.9%	41.3
Compact and Small Station Wagon	1,288,133	33.9	1,320,900	2.5%	40.4
Midsize Car and Station Wagon	1,927,009	33.4	1,984,811	3.0%	39.6
Large Car	332,307	29.0	355,533	6.8%	34.8
Prestige SUV	1,377,565	21.1	1,486,070	7.6%	25.2
Small SUV	1,351,091	23.1	1,467,415	8.3%	27.6
Midsize SUV	285,355	25.9	298,712	4.6%	30.5
Large SUV	1,305,509	27.6	1,343,232	2.8%	32.9
Minivan	719,529	24.9	763,159	5.9%	29.8
Cargo / large passenger van	33,384	19.2	36,071	7.7%	22.9
Cargo Pickup Small	364,995	23.5	374,321	2.5%	28.0
Cargo Pickup Standard	1,377,290	19.7	1,494,324	8.2%	23.6
Ultra Prestige	127,672	20.1	136,388	6.6%	24.1
Total/Average	13,850,480	26.2	14,548,836	4.9%	31.2

For further testing of sensitivity of the results, we varied the payback period, the discount rate, the elasticities, and the start values.

5.1 Payback period

This scenario models the effects of changing the payback period from 1 to 7 years of future fuel savings taken into consideration by vehicle buyers, for the same scenario of a 20 percent increase in fuel economy. Greene (2010) finds a wide range of values in the literature for the willingness of consumers to pay for fuel economy; the payback period is thus a source of uncertainty. In estimating the effects of MY 2012-16 vehicle fuel economy/GHG standards on vehicles, EPA used a 5-year payback period, "which is consistent with the length of a typical new light-duty vehicle loan" (U.S. Environmental Protection Agency and Department of Transportation 2010, p. 25517). The 5-year payback period scenario is the same as the default scenario discussed above.

Total sales increase by approximately 100,000 vehicles, or less than 1 percent, for every year increase in the payback period. Figure 2 shows that shorter payback periods result in less change in market shares, because the change in net price is much smaller when the vehicle buyer ignores more of the future fuel savings. Changes in sales mix, then, become more important as vehicle buyers put more weight on future fuel savings in their purchase decisions.

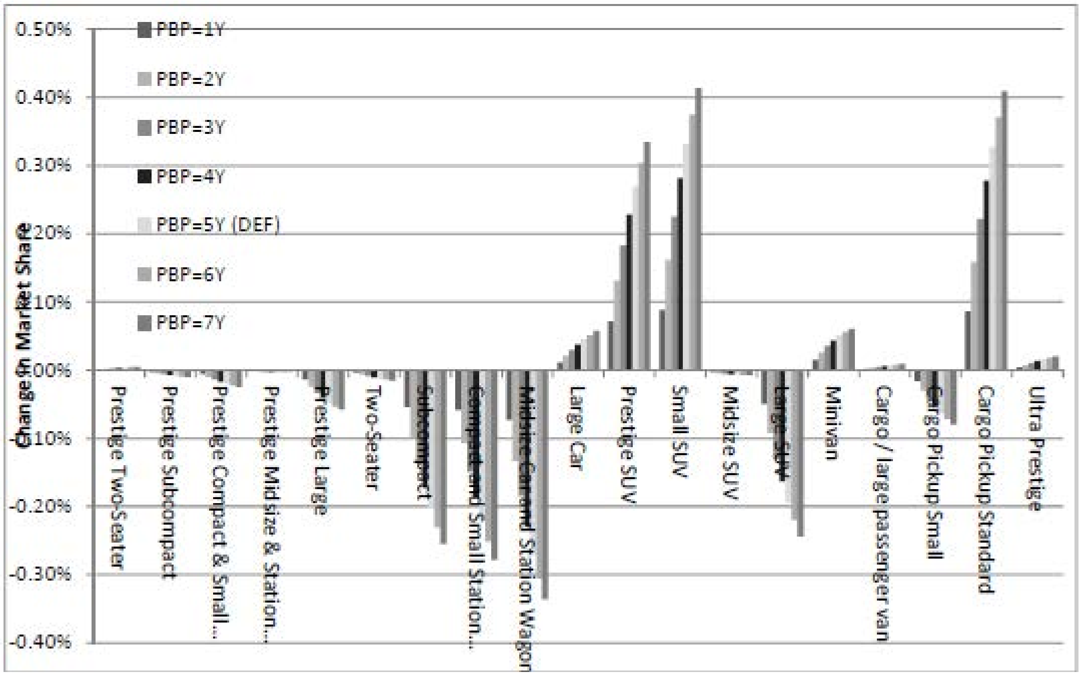

Figure 2: Effects on Market Shares of Changing Payback Period
Note: PBP = payback period, the number of years of fuel savings taken into consideration by vehicle buyers.

5.2 Discount rate

The discount rate enters the model because future fuel savings come to consumers over time; the savings in future years are discounted in the calculation, with larger discount rates reducing the effect of future fuel savings in the net price. As a result, varying the discount rate provides results very similar to those for varying payback period. (The 3 percent discount rate is the same as the default scenario.) Figure 3 shows the effects of changing the discount rate on market shares. It shows the same pattern as the payback period results, with vehicle classes showing the same patterns of gains and losses, because it has similar effects on net price. Because the period facing the discount rate is small (5 years), the magnitude of the effect is small. Varying the discount rate between 2 and 10 percent led to a change in total sales of

111,000 vehicles, less than 1 percent, and about the same amount as changing the payback period by only 1 year.

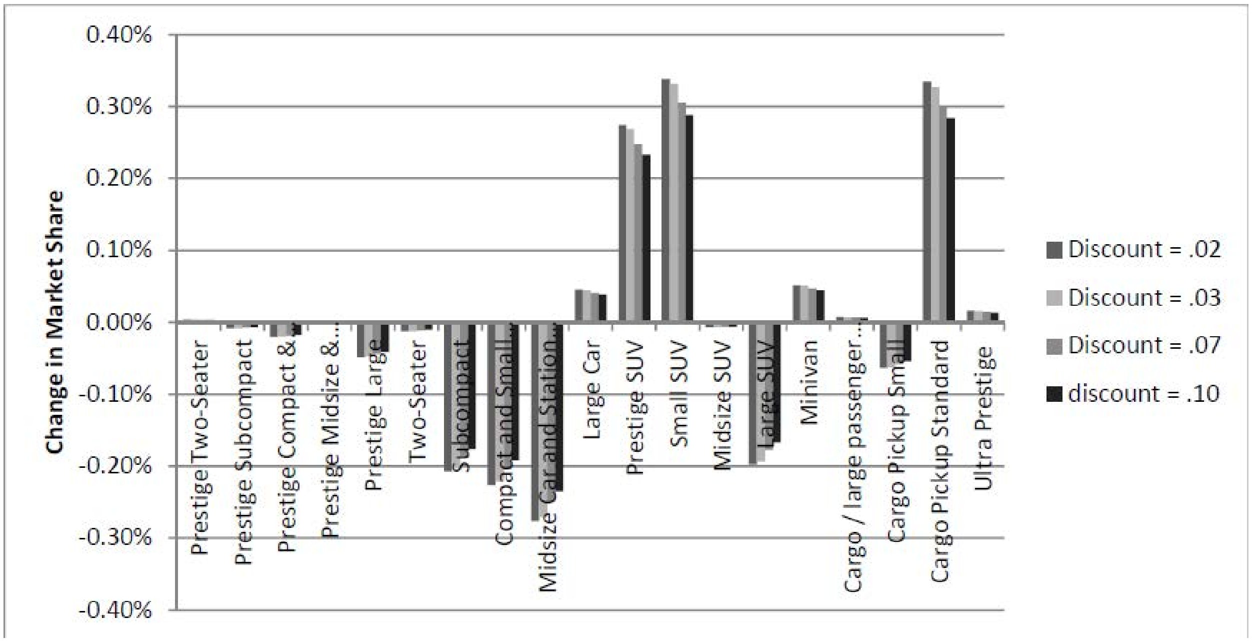

Figure 3: Effects on Market Shares of Different Discount Rates

5.3 Elasticities

To test the effects of the chosen elasticities, we multiplied all the elasticities in Table 2 by 1.5, to see the sensitivity of the analysis to the elasticity values. In addition, we ran scenarios where we multiplied the elasticity of only one class of vehicles by 1.5, to see the effects of changing the elasticity for only one class. The results are in Figure 4.

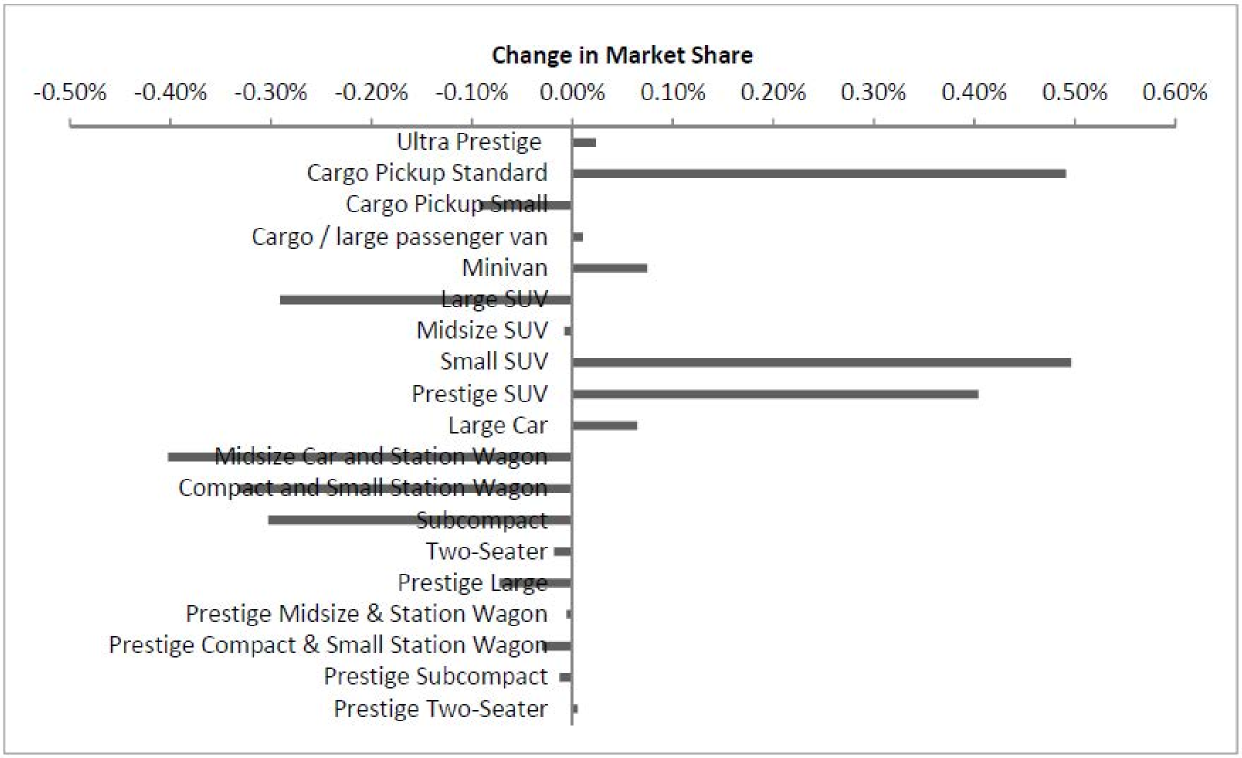

Figure 4: Effects on Market Shares of Multiplying All Elasticities by 1.5

The model predicts that total sales increase by about 1 million vehicles – about 7 percent -- with higher elasticities, compared to the increase of 700,000 vehicles (5 percent) with the baseline elasticities, for the same changes in net prices. With the higher elasticities, vehicle buyers are more responsive to the effective reduction in vehicle prices due to the improvement in fuel economy. Changes in market shares generally exhibit the same pattern of shifts to relatively inefficient vehicles.

When the elasticity of only one class is changed, sales increase by about 700,000 vehicles, almost exactly the amount that vehicle sales changed with the default elasticities, regardless of which vehicle class's elasticity is changed. On average, sales in the class whose elasticity changed increased by about 5 percent, about the same as the increase in sales for the other classes. Market shares also change very little: with average market share per class about 5.3 percent, the maximum change in market share for any scenario was 0.5 percent. This result suggests that the model is not especially sensitive to elasticities in the bottom nest.

5.4 Start Values

To conduct policy projections, the model would need to be calibrated to a baseline fleet in a future year. Because these future fleets in the absence of standards are not known, they provide an additional source of uncertainty for the model. For that reason, we experimented with changing the initial fleet: initially, by multiplying all sales by 1.5; and next, by multiplying initial sales of each individual class by 1.5, holding other class sales constant. We then applied the standard policy scenario of a 20 percent improvement in fuel economy.

The sales response to a change in the fleet size is just about proportional: just as the initial sales increased 4.9 percent in response to the changes in fuel economy, sales with the artificially larger fleet increased 4.9 percent.[10] When sales for individual classes were increased to 150 percent, the sales of that class, and of the remaining classes, increased by about the same proportion as in either the baseline case or when all classes had baseline sales 150 percent higher (see Table 6).

Table 6: Effects of 150 Percent Increases in Baseline Fleet on Changes in Predicted Sales

	Percent Change from 2008 sales due to 20% fuel economy increase	Percent Change from 150% of 2008 sales due to 20% fuel economy increase	Average Percent Change in Own Class Sales when Initial Own Class Sales = 150% of 2008	Percent Change in Named Class's Sales averaged over 18 cases of another class's initial sales 150% of 2008 sales
Prestige Two-Seater	5.59%	5.58%	5.56%	5.76%

[10] Cutting the baseline fleet size in half also produced a 4.9 percent increase in sales due to the policy.

Prestige Subcompact	4.54%	4.53%	4.63%	4.65%
Prestige Compact and Small Station Wagon	4.22%	4.21%	4.30%	4.31%
Prestige Midsize Car and Station Wagon	4.85%	4.84%	4.90%	4.98%
Prestige Large	4.25%	4.24%	4.36%	4.34%
Two-Seater	2.26%	2.25%	2.28%	2.29%
Subcompact	1.92%	1.91%	2.07%	1.93%
Compact and Small Station Wagon	2.51%	2.50%	2.66%	2.54%
Midsize Car and Station Wagon	2.96%	2.94%	3.10%	2.99%
Large Car	6.75%	6.74%	6.87%	7.00%
Prestige SUV	7.58%	7.57%	7.60%	7.89%
Small SUV	8.25%	8.24%	8.13%	8.64%
Midsize SUV	4.57%	4.56%	4.71%	4.68%
Large SUV	2.85%	2.84%	3.23%	2.87%
Minivan	5.89%	5.87%	6.04%	6.06%
Cargo / large passenger van	7.74%	7.73%	8.05%	8.05%
Cargo Pickup Small	2.52%	2.51%	2.81%	2.54%
Cargo Pickup Standard	8.15%	8.14%	8.18%	8.52%
Ultra Prestige	6.60%	6.59%	6.74%	6.83%

This effect is due to the Independence of Irrelevant Alternatives (IIA) characteristic of logit. As mentioned above, this proportionate shifting occurs because, within a nest, the ratio of any two market shares to each other is a constant (Train 2009, Chapters 3, 4); the market shares must change by the same proportion. Across nests, the condition does not hold, and thus the proportions are not exactly constant. Thus, even if the baseline fleet for a policy scenario is inaccurate, the percent change in vehicle sales that it predicts appears to be insensitive to any errors. This finding suggests that percentage increases from the model may be a way to present results that is less sensitive to initial values than presenting sales estimates.

In sum, the sensitivity analyses suggest that the model is not very sensitive to changes in key parameters. Reducing the net prices of vehicles increases sales, with a tendency for sales to move toward those with greater reductions. In addition, changing the default parameters in the model within a reasonable sensitivity range does not have dramatic effect on model outputs. Finally, the percent changes in sales from the model seem to be fairly insensitive to variation in the baseline fleet. These findings may be considered comforting, because all these parameters are subject to a fair degree of uncertainty.

6 Predicting 2010 based on 2008

The sensitivity analysis provides insights into how the model functions. Further validation would come from testing the model's results against actual market outcomes. As straightforward as this goal seems, this is a significant challenge. As has been discussed, this model is meant for a comparative static analysis: if fuel economy and price change, with everything else held constant, how do sales change? In practice, though, holding everything else constant is impossible. In an ideal experimental scenario, a randomized part of the country would face vehicles with the baseline prices and fuel economies, and the rest of the country would face the new prices and fuel economies, at the same time; but that scenario cannot happen when standards apply equally across the country. If, instead, comparisons are made across time, then a number of factors that affect vehicle sales are unlikely to be constant: economic conditions, demographic characteristics, and vehicle characteristics other than fuel economy and price may change. The model is not built to address these questions. As discussed above, the model was not built to be a forecasting model; for the analyses of the effects of potential GHG/fuel economy standards, a comparative statics exercise is appropriate, as long as the model results reflect the response of vehicle markets. This comparative statics framework nevertheless raises significant questions about how to conduct model validation.

With these misgivings as background, we nevertheless compare the model's predictions due to changes in fuel and vehicle prices and changes in fuel economy in MY 2010, relative to MY 2008 vehicles, to those that occurred during MY 2010. As noted above, these model years were used as baseline datasets for the EPA/NHTSA GHG/fuel economy standards for MYs 2017-25 and were thus readily accessible.

The approach is to calibrate the model to MY 2008 vehicle sales, as was done for the sensitivity analyses; provide the model with changes in each vehicle's fuel economy and price between MY 2008 and 2010; and use those changes to predict sales in MY 2010. Those predictions are then compared to actual sales in 2010. Though the Great Recession clearly had a significant effect on the vehicle market during this time -- as Table 3 shows, sales dropped by about 20% -- it is also a period of changes in vehicle characteristics that could be reflected in the modeling results. This period thus provides an opportunity for an initial review of the model's ability to predict changes in the vehicle fleet.

As discussed earlier, we use two datasets: sales-weighted aggregation of different trims into a single model (aggregation by vehicle); and sales-weighted aggregation by manufacturer of all vehicles in a class (aggregation by class). The net change in vehicle price in the model is the change in the vehicle's purchase price, plus some part of the expected lifetime fuel consumption of the vehicle. Expected future fuel consumption used in the model is based on a vehicle's fuel economy, vehicle miles traveled for the specified payback period, fuel prices taken from the Energy Information Administration's Annual Energy Outlook, a 3% discount rate, and a 5 year payback period. As discussed above, the choice of discount rate in the model has a very small effect on the results; the choice of payback period has a small but somewhat larger effect.

7 Results

Table 7 and Table 8 provide an overview of results for the two methods of aggregation. Note that the “actual” market results in the tables omit the vehicles that were not matched between the model years (2 – 9% of all vehicle sales), and thus were excluded from the modeling exercise. This approach assesses the model using all vehicles included in the modeling exercise, rather than the entire population of vehicles. Because each aggregated dataset included a slightly different set of vehicles, the “actual” results are not the same when aggregating by vehicle compared to aggregating by class, as shown in Table 3.

Table 7: Predicted vs. Actual Results with Aggregation by Vehicle

	MY 2008 Actual	MY 2010 Actual	MY 2010 Predicted	MY 2010 Actual – MY 2008 Actual	MY 2010 Actual – MY 2010 Predicted
Total Sales	12,976,766	10,199,188	13,470,888	-2,777,578	-3,271,700
Weighted Avg. Fuel Economy	26.3	28.3	27	2.0	1.3
Share passenger	0.857	0.868	0.849	0.012	0.019
Share cargo	0.134	0.127	0.141	-0.008	-0.014
Share ultra-prestige	0.009	0.005	0.010	-0.004	-0.005
Sum of Absolute Deviations				0.441	0.591
Sum of Squared Deviations				0.0028	0.0051

Table 8: Predicted vs. Results with Aggregation by Class

	MY 2008 Actual	MY 2010 Actual	MY 2010 Predicted	MY 2010 Actual – MY 2008 Actual	MY 2010 Actual – MY 2010 Predicted
Total Sales	13,573,775	10,648,872	14,035,057	-2,924,903	-3,386,185

Weighted Avg. Fuel Economy	26.3	27.5	26.9	-26.3	-26.9
Share passenger	0.860	0.872	0.862	0.012	0.010
Share cargo	0.131	0.121	0.129	-0.009	-0.008
Share ultra-prestige	0.009	0.007	0.009	-0.003	-0.002
Sum of Absolute Deviations				0.4354	0.5769
Sum of Squared Deviations				0.0043	0.0101

Both methods do poorly in predicting total vehicle sales. This result is not a surprise, given the model years studied and the model's function. As discussed above, the model is not designed to predict future vehicle sales based on future changes; instead, it is intended for comparisons within a model-year of vehicles with and without fuel-saving technologies. Sales in MY 2010 were heavily affected by the Great Recession, which the model, calibrated to MY 2008, would not take into account. Both forms of aggregation predict increases in vehicle sales, a result that must be due to decreases in effective prices (price plus a portion of future fuel consumption) between the two years.

Both forms of aggregation correctly predict increases in fuel economy resulting from the change in sales mix, though the actual increase in fuel economy is greater than that predicted in either form of aggregation. This effect may, again, perhaps be due to the influence of the Great Recession: people may have switched to less expensive vehicles, which may tend to be more fuel-efficient than more expensive vehicles. There may be other explanations for this result as well. Perhaps, for instance, people accounted more for future fuel consumption in their purchase decisions than these model runs allowed.

Although the model did not correctly estimate vehicle sales, perhaps it does better in forecasting consumer shifts across vehicle classes in response to changes in price and fuel economy. We thus calculate, for both datasets, the difference between predicted and actual market shares at the vehicle level. In addition, as in Haaf et al., we calculate the difference between actual market shares for MY 2010 and actual market shares for MY 2008, as a naïve, no-information alternative; one simple test of the value of the model is whether it out-performs an alternative of assuming no change in market shares. To compare the two forecasts, we calculate both the summed absolute deviation and the sum of squared deviations from actual MY 2010 market shares for both the modeling results and the MY 2008 actual results.

Table 7 and Table 8 provide results for shifts between passenger cars, cargo vehicles, and ultra-prestige vehicles. At this very aggregated level, actual shifts are small – about 1 percent from cargo vehicles to passenger vehicles. This very small shift may well be due to the way that vehicles are classified in this model. Many vehicle classes that may be legally defined as trucks, such as SUVs, are here considered to be passenger vehicles, because people use them that way. With over 85 percent of vehicles in the passenger category, most shifts are likely to stay within that category, rather than move across categories.

The two methods of aggregation produced opposite results directionally from the model for the shares of passenger, cargo, and ultra-prestige vehicles: aggregation by vehicle implied a switch from passenger vehicles to cargo vehicles, with aggregation by class showing what actually happened, a relative increase in passenger vehicles. These shifts are small: the actual full market share in passenger vehicles went from about 86% to 88% (see Table 4), though either form of aggregation used a slightly smaller share of passenger vehicles. Both aggregations may have omitted slightly more passenger vehicles than cargo or ultra-prestige vehicles, perhaps reflecting a greater tendency of passenger vehicles to be redesigned in ways that make them hard to link across years.

Table 9 and Table 10 provide the results of these comparisons for the 19 vehicle classes in the model. For both datasets, using MY 2008 market shares to forecast MY 2010 market shares out-performs the vehicle choice model when deviations are measured at the class level. This result is similar to those of Haaf et al., who found that using static market shares out-performed attribute-based models when predicting one year ahead.[11]

[11] Haaf et al. also found that attribute-based models could do better than the static market shares model for a four-year-ahead forecast. They find generally that models with more attributes forecast better than models with fewer or no attributes.

Table 9: Class Shifts for Aggregation by Vehicle

Market Shares by Vehicle Class	MY 2008 Actual	MY 2010 Actual	MY 2010 Predicted	MY 2010 Predicted High Midsize Elast	MY 2010 Actual – MY 2008 Actual	MY 2010 Actual – MY 2010 Predicted	MY 2010 Actual – MY 2010 Pred High Mid Elast
Prestige Two-Seater	0.0056	0.0031	0.0052	0.0051	-0.0025	-0.0021	-0.0020
Prestige Subcompact	0.0200	0.0111	0.0193	0.0190	-0.0089	-0.0082	-0.0079
Prestige Compact & Small Station Wagon	0.0293	0.0327	0.0321	0.0316	0.0034	0.0005	0.0011
Prestige Midsize & Station Wagon	0.0378	0.0403	0.0353	0.0347	0.0026	0.0050	0.0056
Prestige Large	0.0671	0.0552	0.0618	0.0607	-0.0119	-0.0066	-0.0055
Two-Seater	0.0029	0.0010	0.0028	0.0027	-0.0019	-0.0018	-0.0017
Subcompact	0.0750	0.0617	0.0858	0.0731	-0.0133	-0.0241	-0.0114
Compact and Small Station Wagon	0.0916	0.1183	0.0939	0.0799	0.0267	0.0245	0.0384
Midsize Car and Station Wagon	0.1388	0.1830	0.1331	0.1735	0.0442	0.0499	0.0095
Large Car	0.0314	0.0223	0.0213	0.0182	-0.0091	0.0010	0.0041
Prestige SUV	0.0957	0.0800	0.0932	0.0916	-0.0157	-0.0132	-0.0116
Small SUV	0.0919	0.0879	0.0997	0.0980	-0.0040	-0.0118	-0.0101
Midsize SUV	0.0220	0.0171	0.0165	0.0162	-0.0049	0.0006	0.0009
Large SUV	0.1004	0.1064	0.0995	0.0978	0.0060	0.0069	0.0086
Minivan	0.0472	0.0481	0.0497	0.0489	0.0009	-0.0016	-0.0008
Cargo / large passenger van	0.0010	0.0017	0.0008	0.0008	0.0008	0.0009	0.0009
Cargo Pickup Small	0.0280	0.0244	0.0227	0.0224	-0.0036	0.0017	0.0019
Cargo Pickup Standard	0.1053	0.1005	0.1174	0.1160	-0.0049	-0.0169	-0.0155
Ultra Prestige	0.0090	0.0052	0.0099	0.0098	-0.0038	-0.0047	-0.0046
Sum of Absolute Deviations					0.1690	0.1820	0.1423
Sum of Squared Deviations					0.0036	0.0045	0.0024

Table 10: Class Shifts for Aggregation by Class

Market Shares by Vehicle Class	MY 2008 Actual	MY 2010 Actual	MY 2010 Predicted	MY 2010 Predicted High Midsize Elast	MY 2010 Actual – MY 2008 Actual	MY 2010 Actual – MY 2010 Predicted	MY 2010 Actual – MY 2010 Pred High Mid Elast
Prestige Two-Seater	0.0051	0.0026	0.0047	0.0047	-0.0025	-0.0021	-0.0021
Prestige Subcompact	0.0220	0.0179	0.0322	0.0320	-0.0041	-0.0143	-0.0141
Prestige Compact & Small Station Wagon	0.0276	0.0329	0.0293	0.0291	0.0053	0.0036	0.0038
Prestige Midsize & Station Wagon	0.0426	0.0313	0.0384	0.0382	-0.0112	-0.0071	-0.0069
Prestige Large	0.0725	0.0288	0.0536	0.0534	-0.0437	-0.0249	-0.0246
Two-Seater	0.0028	0.0010	0.0027	0.0027	-0.0018	-0.0018	-0.0018
Subcompact	0.0662	0.0435	0.0868	0.0822	-0.0226	-0.0432	-0.0387
Compact and Small Station Wagon	0.0949	0.1645	0.1080	0.1023	0.0696	0.0565	0.0622
Midsize Car and Station Wagon	0.1420	0.1790	0.1117	0.1261	0.0371	0.0674	0.0530
Large Car	0.0245	0.0204	0.0148	0.0141	-0.0041	0.0055	0.0063
Prestige SUV	0.0984	0.0948	0.1033	0.1027	-0.0036	-0.0085	-0.0079
Small SUV	0.0975	0.0987	0.1254	0.1247	0.0012	-0.0267	-0.0260
Midsize SUV	0.0210	0.0111	0.0119	0.0118	-0.0099	-0.0008	-0.0007
Large SUV	0.0962	0.0957	0.0901	0.0896	-0.0004	0.0056	0.0061
Minivan	0.0467	0.0496	0.0490	0.0488	0.0029	0.0006	0.0009
Cargo / large passenger van	0.0025	0.0016	0.0022	0.0022	-0.0008	-0.0006	-0.0006
Cargo Pickup Small	0.0268	0.0233	0.0239	0.0238	-0.0034	-0.0005	-0.0005
Cargo Pickup Standard	0.1015	0.0964	0.1029	0.1026	-0.0051	-0.0065	-0.0062
Ultra Prestige	0.0094	0.0067	0.0090	0.0090	-0.0027	-0.0023	-0.0023
Sum of Absolute Deviations					0.2322	0.2786	0.2645
Sum of Squared Deviations					0.0090	0.0114	0.0099

In predicting shares of the 19 vehicle classes included in the vehicle choice model, aggregation by class correctly estimated the direction of shifts in more cases (14 out of 19 classes) than did aggregation by vehicle (10 out of 19 classes) (see Table 9 and Table 10). Most of the shifts in market shares are small, though: in most cases (13 for aggregation by class, 14 for aggregation by vehicle), the predicted market share is within 1 percent of the actual market share. With mostly small changes in market shares, it may be difficult to distinguish the quality of modeling performance from a general tendency for market shares not to change very much.

For both aggregations, the largest class is Midsize Cars and Station Wagons. This class experienced a relatively large shift in shares between 2008 and 2010, from about 14% to 18-19%. Both forms of aggregation not only missed the magnitude of this shift, but even missed the direction. It is not possible to say from which classes people switched (other than the obvious point that people generally switched from classes where shares went down). The relatively inaccurate performance for this large class suggests that it could be useful to experiment with adjusting the demand elasticity for it, though results from the sensitivity analyses suggest that a large elasticity change would be necessary to improve the results substantially.

Table 9 and Table 10 also include results from running the model with an elasticity of -15 – three times the default value – for the Midsize class. For Aggregation by Vehicle, this change leads to an increase in market share for the Midsize class almost as large as occurred, with a reduction in overall deviations as well; for Aggregation by Class, market share increases much less, and using MY 2008 market shares is still the better forecast. This exercise suggests that it is possible to improve the model's match to actual results, but it may match actual results only for MY 2010. Additional data and further testing would be needed to evaluate whether the model's forecasting ability is improved with the revised parameters.

In sum, it is difficult to devise a test of the model, based on its design. As discussed previously, the model is intended to consider the vehicle fleet with and without standards, not the fleet response to changes in social and economic conditions. The ideal test would require having two fleets of otherwise identical vehicles – one with fuel-saving technologies, one without – available for sale at the same time, because that scenario is the one the model is designed to assess. Using data from different model years clearly does not meet this ideal.

It is not a surprise that the model did not predict the reduction in sales due to the recession. It is perhaps a bit more dismaying that it does not show a strong ability to predict changes in market shares, perhaps due to missing changes in tastes or income effects between the two years. Again, though, the model was not designed to consider how the recession may have affected those factors. Given that most market shares, and changes in market shares, are small, it may be difficult to identify those changes even under more consistent conditions. Indeed, the results suggest that holding market shares constant from the initial year may provide better estimates than using the model. Additional work, potentially with additional model years of data, or development of new methods for model validation in the absence of a counter-factual, may be needed to understand better the ability of the EPA model to estimate changes in vehicle purchases associated with changes in vehicle fuel economy.

8 Conclusion

Consumer vehicle choice models are commonly used to simulate the effects of counter-factual situations; they have been tested against actual market outcomes much less frequently. In the few cases where models with forecasting ability have been tested against market outcomes, results are still not very strong, especially for market share predictions. Perhaps innovations in vehicles or changes in consumer preferences in response to advertising or market conditions, or even just the difficulties in properly specifying consumers' preferences, lead to limitations in models' predictive abilities.

This paper adds to, and seeks to encourage, that literature by examining the performance of a model developed for the U.S. Environmental Protection Agency to estimate responses to changes in vehicle prices and fuel economy. The ability to adjust parameters easily allows for sensitivity analysis with alternative assumptions of model parameters. The sensitivity analyses suggest that moderate changes in the default parameters and baseline fleet have small effects on the model outputs. Given the uncertainties associated with many of these parameters, this finding suggests some robustness of modeling results to those uncertainties. The test of the model against actual market outcomes suggests that the model is not suitable for forecasting changes in the vehicle fleet when social and economic conditions are also changing. Because the model was not designed to forecast such changes, this result is expected. It is nevertheless not encouraging for model validation that assuming the market shares of the base fleet had less forecast error than using the model.

Perhaps the major lesson is that conducting a validation exercise can be a significant challenge, and perhaps other approaches may be needed to validate a model designed for policy simulation rather than forecasting. First, as already mentioned, there is no obvious way to test the model for the purpose for which it was designed, because only one vehicle fleet exists in the U.S. in a year; no counter-factual exists. Vehicle choice models that incorporate demographic factors and vehicle attributes may be better suited to testing across time; on the other hand, when they are ultimately used for simulation purposes, such models require projections of those demographic factors and vehicle attributes, which may not be of great reliability. Across time, any model has to face the fact that vehicles, even manufacturers, enter and exit the market. Whitefoot et al. (2013) seek to endogenize manufacturer and consumer decisions simultaneously; whether such efforts will reflect actual market movements is yet to be seen.

The results presented here suggest that further work is desirable. For instance, it would be valuable to analyze additional years of data. Do predictions of responses to vehicles in future model years follow the same pattern as in MY 2010? Or might the model predict better for non-recession years, or worse for years further in the future? If adjustments to model parameters improve forecasts for MY 2010 market shares, would those adjustments work as well for other years? For other researchers, this paper aims to encourage further work on validation of other models, both in development of methods and in application of those methods. We hope that this paper stimulates more research on the ability of consumer vehicle choice models to predict actual market changes.

9 References

Allcott, H. (2013). "The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market." *American Economic Journal: Economic Policy* 5(3): 30-66.

Austin, D., and T. Dinan (2005). "Clearing the Air: The Costs and Consequences of Higher CAFE Standards and Increased Gasoline Taxes." *Journal of Environmental Economics and Management* 50: 562-582.

Bento, A.M., S. Li, and K. Roth (2012). "Is There an Energy Paradox in Fuel Economy? A Note on the Role of Consumer Heterogeneity and Sorting Bias," *Economics Letters* 115: 44-48.

Berry, S., J. Levinsohn, and A. Pakes (1995). "Automobile Prices in Market Equilibrium," *Econometrica* 63(4): 841-940.

Brownstone, D., D. Bunch, T. Golob, and W. Ren (1996). "A Transactions Choice Model for Forecasting Demand for Alternative-Fuel Vehicles." *Research in Transportation Economics*, 4: 87-129.

Energy Information Administration (2008). "Annual Energy Outlook 2008." DOE/EIA-0383, http://www.eia.gov/oiaf/aeo/pdf/0383%282008%29.pdf .

Goldberg, P. (1998). "The Effects of the Corporate Average Fuel Efficiency Standards in the U.S." *Journal of Industrial Economics* 46(1): 1-33.

Greene, D. (2009). "Feebates, Footprints and Highway Safety," *Transportation Research Part D* 14: 375-384.

Greene, D. (2010). "How Consumers Value Fuel Economy: A Literature Review." Office of Transportation and Air Quality, U.S. Environmental Protection Agency, EPA-420-R-10-008.

Greene, David L., and Changzheng Liu (2012). "Consumer Vehicle Choice Model Documentation." U.S. Environmental Protection Agency EPA-420-B-12_052, http://www.epa.gov/otaq/climate/documents/420b12052.pdf.

Haaf, C.G., J.J. Michalek, W.R. Morrow, and Y. Liu (2014). "Sensitivity Of Vehicle Market Share Predictions to Discrete Choice Model Specification." *Journal of Mechanical Design* 136, 121402-121402-9.

Helfand, G., and A. Wolverton (2011). "Evaluating the Consumer Response to Fuel Economy: A Review of the Literature." *International Review of Environmental and Resource Economics* 5: 103-146.

Jacobsen, M. (2013). "Evaluating U.S. Fuel Economy Standards in a Model with Producer and Household Heterogeneity." *American Economic Journal: Economic Policy* 5(2): 148-87.

Knittel, C.R., and K. Metaxoglou (2014). "Estimation of Random-Coefficient Demand Models: Two Empiricists' Perspective." *Review of Economics and Statistics* 96(1): 34-59.

Landry, C.E., and J. List (2007). "Using Ex Ante Approaches to Obtain Credible Signals for Value in Contingent Valuation Markets: Evidence from the Field." *American Journal of Agricultural Economics* 89: 420-429.

Larrick, R.P., and J.B. Soll (2008). "The MPG Illusion." *Science* 320(5883): 1593-94.

Pakes, A., S. Berry, and J. Levinsohn (1993). "Applications and Limitations of Some Recent Advances in Empirical Industrial Organization: Price Indexes and the Analysis of

Environmental Change." *American Economic Review Papers and Proceedings* 83(2): 240-246.

Raynaert, Mathias (2014). "Abatement Strategies and the Cost of Environmental Regulation: Emission Standards on the European Car Market." KU Leuven Center for Economic Studies Discussion Paper Series DPS14.31.

Train, K. (2009). *Discrete Choice Methods with Simulation.* Cambridge University Press. http://elsa.berkeley.edu/books/choice2.html

Train, K., and Clifford Winston (2007). "Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers," *International Economic Review* 48: 1469-1496.

U.S. Environmental Protection Agency (2014). "Operational Performance Evaluation of Air Quality Model Simulations." http://www.epa.gov/AMD/Research/Air/operationalEval.html, accessed September 2, 2014.

U.S. Environmental Protection Agency (2012). "Optimization Model for reducing Emissions of Greenhouse gases from Automobiles (OMEGA)." http://www.epa.gov/otaq/climate/models.htm, accessed September 8, 2014.

U.S. Environmental Protection Agency and Department of Transportation (2010). "Light-Duty Vehicle Greenhouse Gas Emissions Standards and Corporate Average Fuel Economy Standards; Final Rule." *Federal Register* 75(88): 25324-25728.

U.S. Environmental Protection Agency and Department of Transportation (2012). "Joint Technical Support Document: Final Rulemaking for 2017-25 Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards."

EPA-420-R-12-901, Chapter 1.
http://www.epa.gov/otaq/climate/documents/420r12901.pdf

Whitefoot, K., M. Fowlie, and S. Skerlos (2013). "Compliance by design: Industry response to efficiency standards." Working paper, http://nature.berkeley.edu/~fowlie/whitefoot_fowlie_skerlos_submit.pdf

Whitefoot, K., and S. Skerlos (2012). "Design Incentives to Increase Vehicle Size Created from the U.S. Footprint-Based Fuel Economy Standards." *Energy Policy* 41: 402-411.0